CLAIMS

We claim:

1. A method of generating an index to provide access to a database where the database includes structured data columns and unstructured data columns, the method comprising the steps of:

identifying one or more unstructured data columns and one or more structured data columns from the database;

generating a combined index of the unstructured and structured data columns, the combined index providing associative access to the database using a query that includes both unstructured conditions and structured conditions.

2. The method as set forth in claim 1 further including generating an inverted index table based on the unstructured data column.

3. The method as set forth in claim 2 further including tokenizing the unstructured data column into tokens where the inverted index table is generated based on the tokens, the tokens being data from the unstructured data column.

4. The method as set forth in claim 3 further including generating a B-tree from the inverted index table to form the combined index, where a first level of branching is based on the tokens and a second level of branching is based on values from the structured data column.

5. The method as set forth in claim 3 wherein the unstructured data column includes text, and where the tokenizing includes parsing the text into individual words that become the tokens.

6. The method as set forth in claim 2 further including adding the structured data column to the inverted index.

7. The method as set forth in claim 1 wherein the unstructured data column includes text.

8. The method as set forth in claim 1 wherein the unstructured data column includes data being one of image data, video data, and audio data.

9. The method as set forth in claim 8 further including:

generating a signature for each of the data of the unstructured data column;

generating an inverted index table based on the signature for each of the data and the structured data column associated with each signature; and

generating a tree index of the inverted index table to form the combined index.

10. A database management system for providing access to a data table, the system comprising:

a user interface for receiving instructions from a user; and

indexing logic that generates an index structure by combining an unstructured data column and a structured data column from the data table,

the index structure allowing data retrieval from the data table based on a query received from the user interface having conditions associated to both the unstructured data column and the structured data column.

11. The database management system as set forth in claim 10 wherein the index structure is a B-tree including branches for data from the unstructured data column and branches for data from the structured data column.

12. The database management system as set forth in claim 10 wherein the unstructured data column includes rows having text data including one or more words, and the structured data column includes corresponding rows having a value.

13. The database management system as set forth in claim 12 wherein the indexing logic further includes:

segmenting logic for segmenting the text data from each of the rows of the unstructured data column into words; and

logic for building an inverted index table from the data table based on the words from the unstructured data column and corresponding values from the structured data column.

14. The database management system as set forth in claim 13 wherein the index structure is a B-tree generated from the inverted index table where the B-tree forms an index structure of the words from the unstructured data column and the corresponding values from the structured data column.

15. The database management system as set forth in claim 10 wherein the unstructured data column includes rows having one of image data, video data and audio data represented by a signature having searchable characteristics.

16. A method of searching a data table having at least a column of structured data and a column of unstructured data, the method comprising the steps of:

converting the unstructured data into a structured-like form;

generating an index structure based on the converted unstructured data and the structured data; and

searching the data table using the index structure in response to a query having conditions from both the unstructured data and the structured data.

17. The method as set forth in claim 16 wherein the unstructured data includes text data, and the converting includes tokenizing the text data into words.

18. The method as set forth in claim 17 wherein the generating includes building an inverted index table for the data table, the inverted index having the words of the text data as one data column and having values from the structured data as another data column.

19. The method as set forth in claim 18 wherein the index structure is generated as a B-tree from the inverted index table having the words as branching conditions and having the values from the structured data as sub-branching conditions associated with the words.

20. The method as set forth in claim 16 wherein the index structure is an indexing tree having the converted unstructured data as branching conditions and having the structured data as branching conditions.

21. The method as set forth in claim 16 wherein the structured data is an unstructured data converted to a structured form.